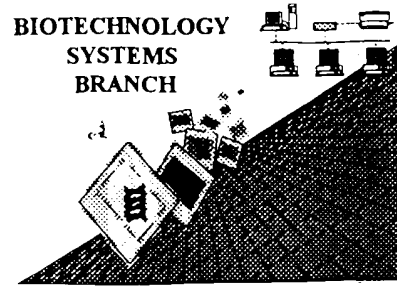


05-90  
0821



## **RAW SEQUENCE LISTING** **ERROR REPORT**

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/923,830  
Source: O/PE  
Date Processed by STIC: 8/16/2001

**THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.**

**PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:**

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,**
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY**

**FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.**

**FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.**

**PATENTIN 2.1 e-mail help: [patin21help@uspto.gov](mailto:patin21help@uspto.gov) or phone 703-306-4119 (R. Wax)**

**PATENTIN 3.0 e-mail help: [patin3help@uspto.gov](mailto:patin3help@uspto.gov) or phone 703-306-4119 (R. Wax)**

**TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 3.0 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:**

### **Checker Version 3.0**

The Checker Version 3.0 application is a state-of-the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 - 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.

Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO). Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

**Checker Version 3.0 can be down loaded from the USPTO website at the following address:**

**<http://www.uspto.gov/web/offices/pac/checker>**

# Raw Sequence Listing Error Summary

ERROR DETECTED	SUGGESTED CORRECTION	SERIAL NUMBER: 09/923,830
ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE		
1 <input checked="" type="checkbox"/> Wrapped Nucleics Wrapped Aminos	The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."	
2 <input type="checkbox"/> Invalid Line Length	The rules require that a line not exceed 72 characters in length. This includes white spaces.	
3 <input type="checkbox"/> Misaligned Amino Numbering	The numbering under each 5 <sup>th</sup> amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.	
4 <input checked="" type="checkbox"/> Non-ASCII	The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.	
5 <input type="checkbox"/> Variable Length	Sequence(s) _____ contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>.<223> section that some may be missing.	
6 <input type="checkbox"/> PatentIn 2.0 "bug"	A "bug" in PatentIn version 2.0 has caused the <220>.<223> section to be missing from amino acid sequences(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>.<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>.<223> sections for Artificial or Unknown sequences.	
7 <input type="checkbox"/> Skipped Sequences (OLD RULES)	Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence: (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading) (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) This sequence is intentionally skipped  Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.	
8 <input type="checkbox"/> Skipped Sequences (NEW RULES)	Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence. <210> sequence id number <400> sequence id number 000	
9 <input checked="" type="checkbox"/> Use of n's or Xaa's (NEW RULES)	Use of n's and/or Xaa's have been detected in the Sequence Listing. Per 1.823 of Sequence Rules, use of <220>.<223> is MANDATORY if n's or Xaa's are present. In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.	
10 <input checked="" type="checkbox"/> Invalid <213> Response	Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>.<223> section is required when <213> response is Unknown or is Artificial Sequence	
11 <input type="checkbox"/> Use of <220>	Sequence(s) _____ missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section. (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)	
12 <input type="checkbox"/> PatentIn 2.0 "bug"	Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.	
13 <input type="checkbox"/> Misuse of n	n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.	

OIPE

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/923,830

DATE: 02/16/2001

TIME: 13:38:20

Input Set : A:\ES.txt

Output Set: N:\CRF3\08162001\I923830.raw

3 <110> APPLICANT: Nienaber, Vicki  
 4 Greer, Jonathan  
 5 Akad-Capatero, Celerino  
 6 Nirbeck, Daniel  
 7 <100> TITLE OF INVENTION: LIGAND SCREENING AND DESIGN BY X-RAY  
 8 CRYSTALLOGRAPHY  
 9 <130> FILE REFERENCE: 6308.US.P1  
 10 <140> CURRENT APPLICATION NUMBER: US/09/923,830  
 11 <141> CURRENT FILING DATE: 2001-08-07  
 12 <150> PRIOR APPLICATION NUMBER: 09/036,184  
 13 <151> PRIOR FILING DATE: 1998-03-06  
 14 <160> NUMBER OF SEQ ID NOS: 14  
 15 <170> SOFTWARE: FastSEQ for Windows Version 3.0

Does Not Compay  
 Corrected Diskette Needed

see item 4  
 on Enr  
 summary  
 sheet  
 pp 1-3

## ERRORED SEQUENCES

20 <210> SEQ ID NO: 1  
 21 <211> LENGTH: 51  
 22 <212> TYPE: DNA  
 23 <213> ORGANISM: Synthetic  
 24 <400> SEQUENCE: 1  
 E--> 26 attaatgtcg actaaggagg tgatctaagc ttaaaatttc agtgtggcca a  
 27 51  
 28 <210> SEQ ID NO: 2  
 29 <211> LENGTH: 57  
 30 <212> TYPE: DNA  
 31 <213> ORGANISM: Synthetic  
 32 <400> SEQUENCE: 2  
 E--> 35 attaataagc ttccagaggg ccaggccatt ctcttccttg gtgtgactcc tgatcca  
 36 57  
 37 <210> SEQ ID NO: 3  
 38 <211> LENGTH: 47  
 39 <212> TYPE: DNA  
 40 <213> ORGANISM: Synthetic  
 41 <400> SEQUENCE: 3  
 E--> 44 attaattgag cagccatccc ggactataca gaccatcgcc ctgccct  
 45 47  
 46 <210> SEQ ID NO: 4  
 47 <211> LENGTH: 46  
 48 <212> TYPE: DNA  
 49 <213> ORGANISM: Synthetic  
 50 <400> SEQUENCE: 4  
 E--> 53 attaatcagc tgctccggat agagatagtc ggtagactgc tctttt  
 54 46  
 55 <210> SEQ ID NO: 5  
 56 <211> LENGTH: 45

(global error)

Unavail - see item 10 on Enr summary  
 sheet

51

see item 1

on Enr

summary

sheet

(global format error)

same error

same

same

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/923,830

DATE: 08/16/01  
TIME: 13:35:20

Input Set : A:\ES.txt

Output Set: N:\CRF3\08162001\I923830.raw

62-210> TYPE: DNA  
62-211> ORGANISM: Synthetic  
62-400> SEQUENCE: 5

E--&gt; 62 attaatcagc tgaaaatgac tgttgtag

*same*

62-210> SEQ ID NO: 6  
62-211> LENGTH: 51

62-212&gt; TYPE: DNA

62-213> ORGANISM: Synthetic  
62-400> SEQUENCE: 6

E--&gt; 71 attaatgtcg actaaggagg tgatctaag ttaaaatttc agtgtggcca a

*same*

71-210> SEQ ID NO: 7  
71-211> LENGTH: 37

71-212&gt; TYPE: DNA

71-213> ORGANISM: Synthetic  
71-400> SEQUENCE: 7

E--&gt; 80 attaatgcta gcctcgagcc accatgagag cctgtgt

*same*

80-210> SEQ ID NO: 8  
80-211> LENGTH: 42

80-212&gt; TYPE: DNA

80-213> ORGANISM: Synthetic  
80-400> SEQUENCE: 8

E--&gt; 89 attaatgcta gcctcgagtc acttggtgtg actgcggatc ca

*same*

89-210> SEQ ID NO: 9  
89-211> LENGTH: 44

89-212&gt; TYPE: DNA

89-213> ORGANISM: Synthetic  
89-400> SEQUENCE: 9

E--&gt; 98 ggtggtgaat tctccccaa taatgccttt ggagtcgctc acga

*same*

98-210> SEQ ID NO: 10  
98-211> LENGTH: 111

98-212&gt; TYPE: DNA

98-213> ORGANISM: Yeast Pichia Pastoris  
98-400> SEQUENCE: 10

E--&gt; 107 atgttctctc caattttgtc cttggaaatt attttagctt tggctacttt gcaatctgtc

107-210&gt;

E--&gt; 109 ttgcgtcagc cagttatctg cactaccgtt ggttcgctg ccgagggatc c

*same*

109-210> SEQ ID NO: 11  
109-211> LENGTH: 22

109-212&gt; TYPE: DNA

109-213> ORGANISM: Synthetic  
109-400> SEQUENCE: 11

E--&gt; 118 gaaacttcca aaagtcgcca ta

118-210&gt;

*same*

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/923,830

DATE: 08/16/01  
TIME: 13:38:20

Input Set : A:\ES.txt

Output Set : N:\CRF3\08162001\I923830.raw

121 &lt;210&gt; SEQ ID NO: 12

122 &lt;211&gt; LENGTH: 92

123 &lt;212&gt; TYPE: DNA

124 <213> ORGANISM: Synthetic

126 &lt;400&gt; SEQUENCE: 12

E--> 127 attaatgaat tctctgagcg gtccgggata cctcggcagc ggaaccaacg gtagtgcaga  
128 .0

E--&gt; 129 taactggctg agcgaagaca gattgcaaag ta

130 .0

132 &lt;210&gt; SEQ ID NO: 13

133 &lt;211&gt; LENGTH: 46

134 &lt;212&gt; TYPE: DNA

135 <213> ORGANISM: Synthetic

137 &lt;400&gt; SEQUENCE: 13

E--> 138 attaatggat ccttggacaa gaggattatt gggggagaat tcacca  
139 .0

141 &lt;210&gt; SEQ ID NO: 14

142 &lt;211&gt; LENGTH: 47

143 &lt;212&gt; TYPE: DNA

144 <213> ORGANISM: Synthetic

146 &lt;400&gt; SEQUENCE: 14

E--> 147 attaatctcg agcggtcctg cacttgggtg gactgcgaat ccagggt  
148 .0

E--&gt; 150 43

E--&gt; 152 6308.us.dl

E--&gt; 156 43

E--&gt; 157 -

E--&gt; 158 43

E--&gt; 160 43

## VERIFICATION SUMMARY

PATENT APPLICATION: US/09/923,830

DATE: 04/07/2001

TIME: 13:55:11

Input Set : A:\ES.txt

Output Set: N:\CRF3\08162001\I923830.raw

L:13 M:270 C: Current Application Number differs, Replaced Current Application No  
 L:13 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
 L:26 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:11 SEQ:1  
 L:35 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:17 SEQ:2  
 L:44 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:47 SEQ:3  
 L:53 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:46 SEQ:4  
 L:62 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:18 SEQ:5  
 L:71 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:11 SEQ:6  
 L:80 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:37 SEQ:7  
 L:89 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:42 SEQ:8  
 L:98 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:44 SEQ:9  
 L:107 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:60 SEQ:10  
 M:254 Repeated in SeqNo=10  
 L:118 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:22 SEQ:11  
 L:127 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:60 SEQ:12  
 M:254 Repeated in SeqNo=12  
 L:138 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:46 SEQ:13  
 L:147 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:47 SEQ:14  
 M:254 Repeated in SeqNo=14  
 L:152 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:6  
 L:152 M:112 C: (48) String data converted to lower case,  
 L:157 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:1  
 L:160 M:252 E: No. of Seq. differs, <211>LENGTH:Input:47 Found:53 SEQ:14